

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2005 年 5 月 26 日 (26.05.2005)

PCT

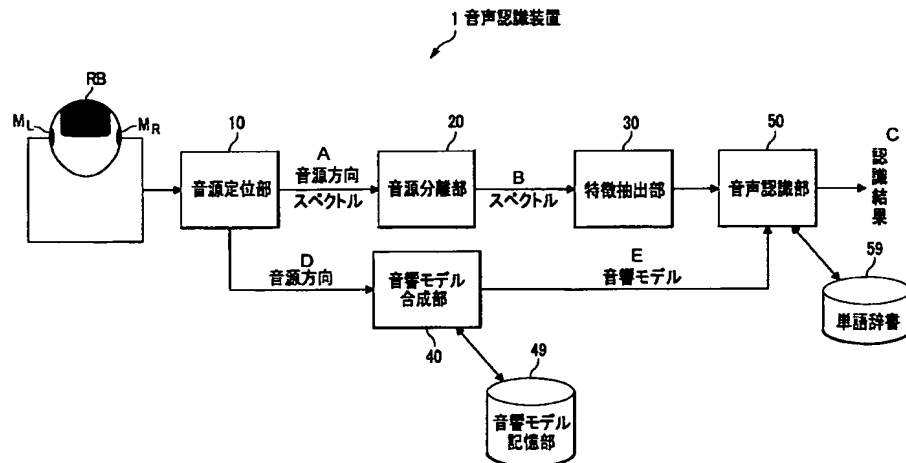
(10) 国際公開番号  
WO 2005/048239 A1

- (51) 国際特許分類<sup>7</sup>: G10L 15/06 (72) 発明者; および  
(21) 国際出願番号: PCT/JP2004/016883 (75) 発明者/出願人 (米国についてののみ): 中臺 一博  
(22) 国際出願日: 2004 年 11 月 12 日 (12.11.2004) (NAKADAI, Kazuhiro) [JP/JP]; 〒3510193 埼玉県和  
(25) 国際出願の言語: 日本語 光市中央 1 丁目 4-1 Saitama (JP). 辻野 広司 (TSU-  
(26) 国際公開の言語: 日本語 JINO, Hiroshi) [JP/JP]; 〒3510193 埼玉県和光市中央  
(30) 優先権データ: 1 丁目 4-1 Saitama (JP). 奥乃 博 (OKUNO, Hiroshi)  
特願 2003-383072 [JP/JP]; 〒6048135 京都府京都市中京区東洞院通三  
2003 年 11 月 12 日 (12.11.2003) JP 条下る三文字町 2 0 5 番地の 3 フォーム東洞院三  
(71) 出願人 (米国を除く全ての指定国について): 本田技研 工業株式会社 (HONDA MOTOR CO., LTD.) [JP/JP]; 山本 俊一 (YAMAMOTO, Shunichi) [JP/JP]; 〒3510188 埼玉県和光市本町 8-1  
〒1078556 東京都港区南青山 2-1-1 Tokyo (JP). (74) 代理人: 磯野 道造 (ISONO, Michizo); 〒1020093 東京  
都千代田区平河町 2 丁目 7 番 4 号 砂防会館別館内  
磯野国際特許商標事務所気付 Tokyo (JP).

[続葉有]

(54) Title: SPEECH RECOGNITION DEVICE

(54) 発明の名称: 音声認識装置



1... SPEECH RECOGNITION DEVICE  
10... SOUND SOURCE LOCALIZATION UNIT  
A... SOUND SOURCE DIRECTION SPECTRUM  
20... SOUND SOURCE ISOLATION UNIT  
B... SPECTRUM  
30... FEATURE EXTRACTION UNIT  
50... SPEECH RECOGNITION UNIT

C... RECOGNITION RESULT  
D... SOUND SOURCE DIRECTION  
40... ACOUSTIC MODEL SYNTHESIS UNIT  
E... ACOUSTIC MODEL  
59... WORD DICTIONARY  
49... ACOUSTIC MODEL STORAGE UNIT

(57) Abstract: There is provided a speech recognition device capable of recognizing a speech with a high accuracy even when a speaker or mobile body having the speech recognition device is moving. The speech recognition device recognizes a speech of a particular speaker HMj from an acoustic signal detected by a plurality of microphones M and converts it into character information. The speech recognition device includes: a sound source localization unit (10) for localizing the sound source direction  $\theta_{HMj}$  of the speaker HMj; a sound source isolation unit (20) for isolating the speech signal of the speaker HMj from the acoustic signal according to the sound source direction  $\theta_{HMj}$ ; an acoustic model storage unit (49) containing a direction-dependent acoustic model  $H(\theta_n)$  corresponding to intermittent plurality of directions; an acoustic model synthesis unit (40) for obtaining an acoustic model of the sound source direction  $\theta_{HMj}$

[続葉有]



(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:  
— 国際調査報告書

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY,

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

according to the direction-dependent acoustic model  $H(\theta_n)$  of the acoustic model storage unit (49) and making the acoustic model storage unit (49) store the result; and a speech recognition unit (50) for using the acoustic model synthesized by the acoustic model synthesis unit (40) so as to perform speech recognition of the audio signal isolated by the sound source isolation unit (20) and convert the speech into character information.

(57) 要約: 話者や、音声認識装置を搭載した移動体が移動しても高い精度で認識可能な音声認識装置を提供する。複数のマイクMが検出した音響信号から、特定の話者HMjの音声を認識して文字情報に変換する音声認識装置である。複数のマイクMが検出した音響信号に基づき、話者HMjの音源方向 $\theta_{HMj}$ を特定する音源定位部10と、音源方向 $\theta_{HMj}$ に基づき、話者HMjの音声信号を音響信号から分離する音源分離部20と、断続的な複数の方向に対応した方向依存音響モデル $H(\theta_n)$ を記憶した音響モデル記憶部49と、音源方向 $\theta_{HMj}$ の音響モデルを、音響モデル記憶部49の方向依存音響モデル $H(\theta_n)$ に基づいて求め、音響モデル記憶部49へ記憶させる音響モデル合成部40と、音響モデル合成部40が合成した音響モデルを使用して、音源分離部20が分離した音声信号の音声認識を行い、文字情報に変換する音声認識部50とを備える。